



Purpose

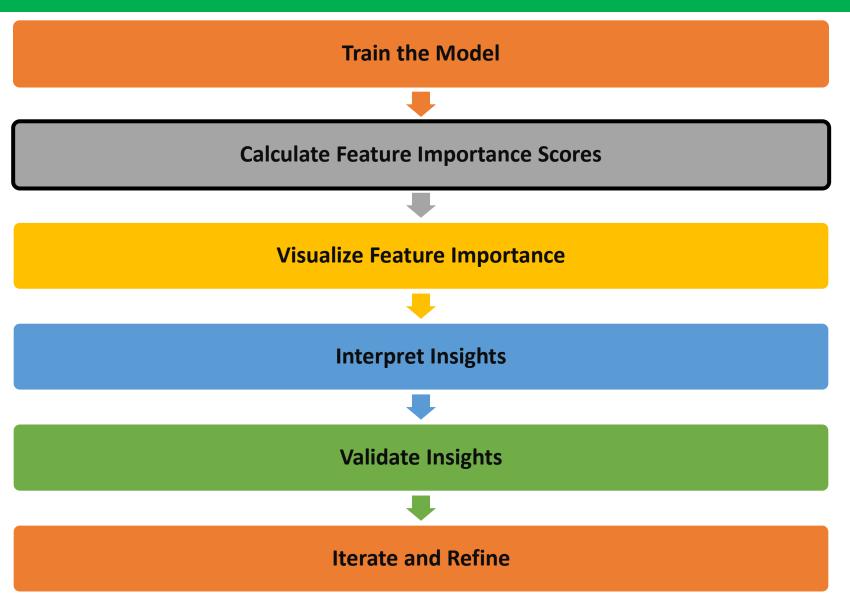
The purpose of the section is to help you learn how to research, select, and develop appropriate algorithms to become a Successful Artificial Intelligence (AI) Engineer

At the end of this lecture, you will learn the following

 How to use feature importance analysis to provide insights into model predictions













Calculate Feature Importance Scores

Decision Trees

How much each feature decreases the impurity in the tree nodes

Random Forests

Averaging the decrease in impurity over all trees in the forest.

Gradient Boosting Machines

Total decrease in loss (e.g., mean squared error) attributed to each feature across all the trees.

Linear Models

Coefficients of the features



Enrichmentor

What is impurity in the tree nodes?

Decision Trees

How much each feature decreases the impurity in the tree nodes (e.g., Gini importance or information gain)

Random Forests

Averaging the decrease in impurity over all trees in the forest

Gradient Boosting Machines

Total decrease in loss (e.g., mean squared error) attributed to each feature across all the trees

Linear Models

Coefficients of the features



Enrichmentor

What is impurity in the tree nodes?

Gini Impurity

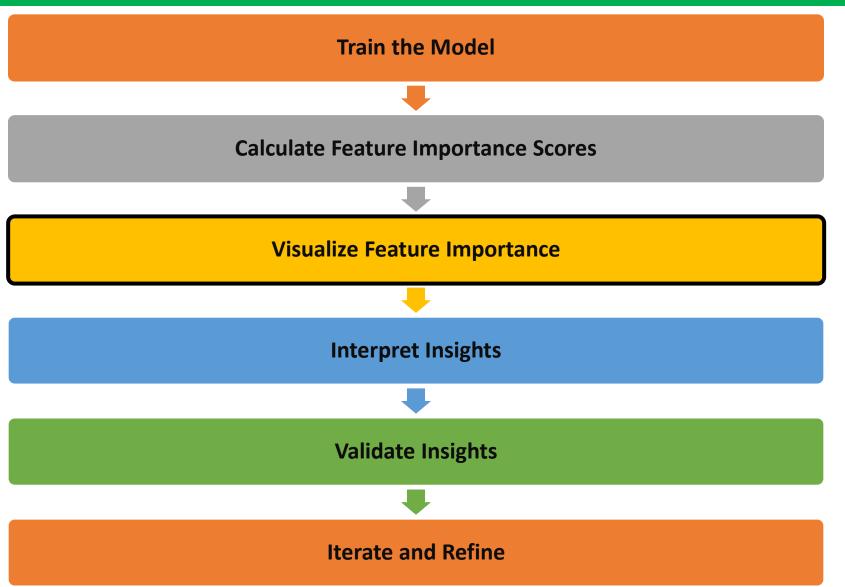
- Measures the probability of incorrectly classifying a randomly chosen element
- Ranges from 0 to 0.5, where 0 indicates that the node is pure and 0.5 indicates maximum impurity

Entropy

- Measures the average amount of information needed to classify a sample in a dataset
- Calculated as the sum of the probabilities of each class label multiplied by the logarithm of that probability
- Ranges from 0 to 1, where 0 indicates a pure node and 1 indicates maximum impurity

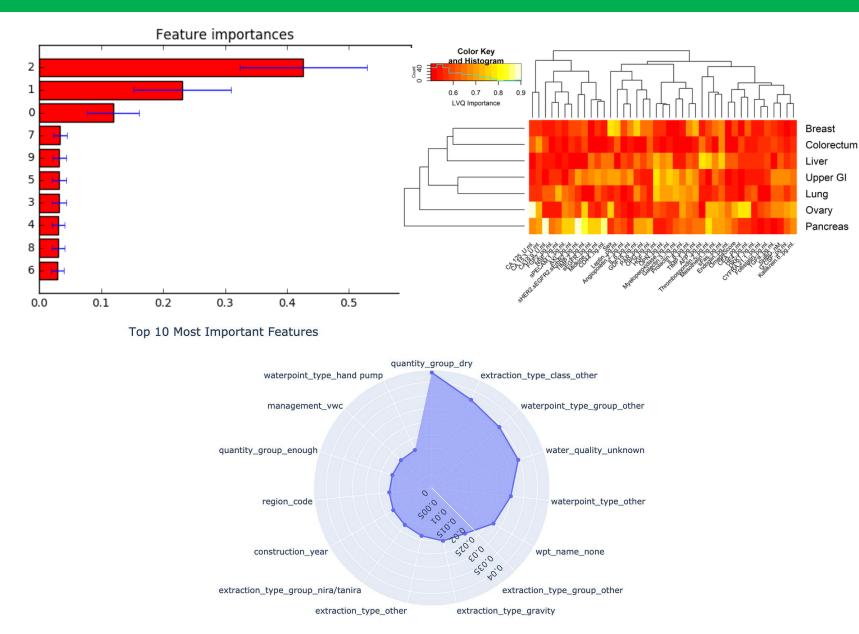










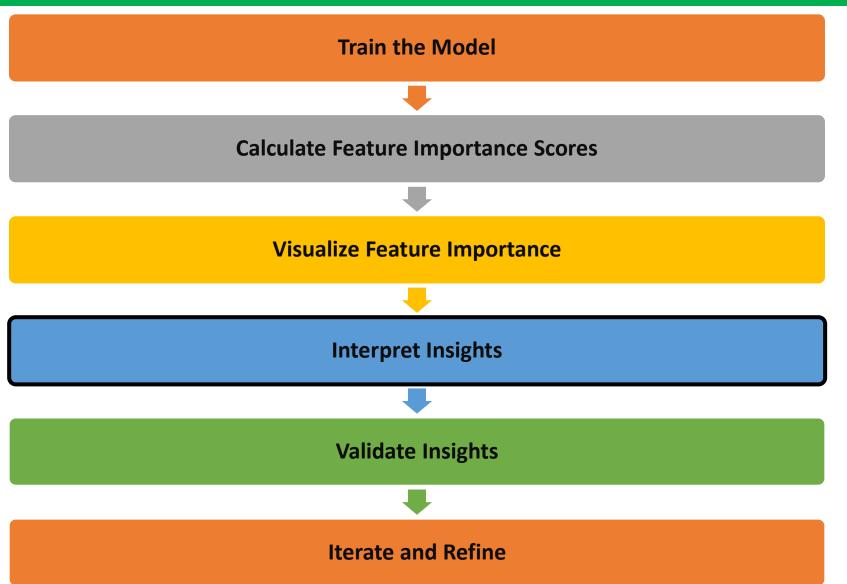






Growing through Excellence over 40 years to become Best in Management









Which features have the highest importance scores?

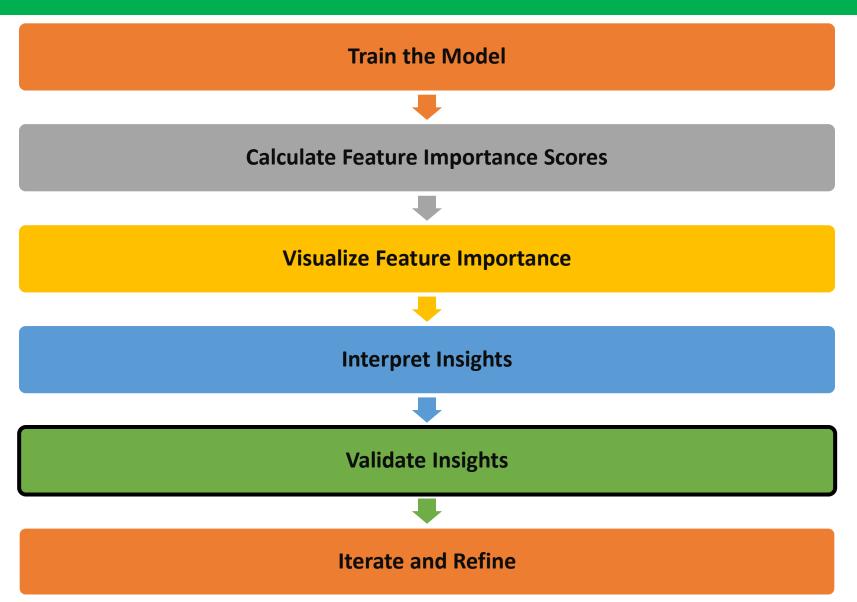
Are there any correlations or interactions between features that influence their importance?

Are there any surprising or unexpected results?









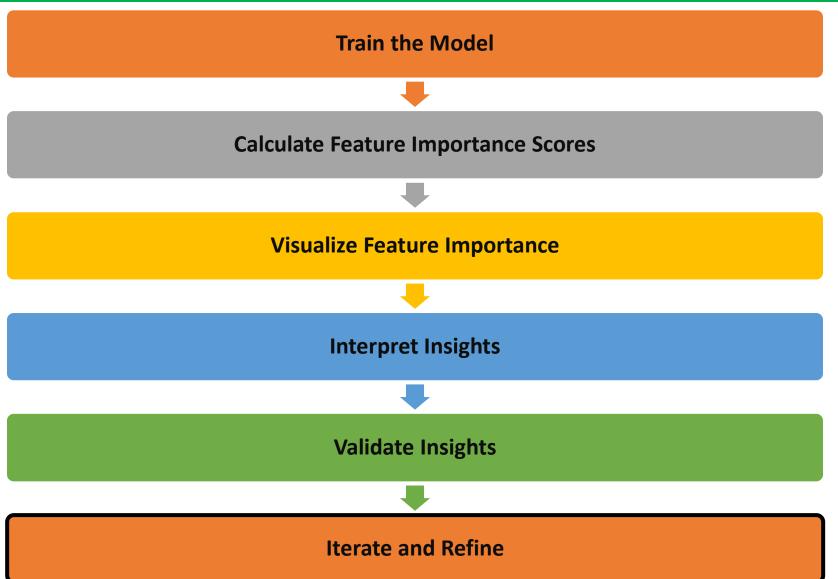




Insights Validation Examining how Conducting further well they align with experiments or domain knowledge analysis Testing the impact of removing certain Conducting features on model sensitivity analyses performance



Enrichmentors







iteratively improve your model

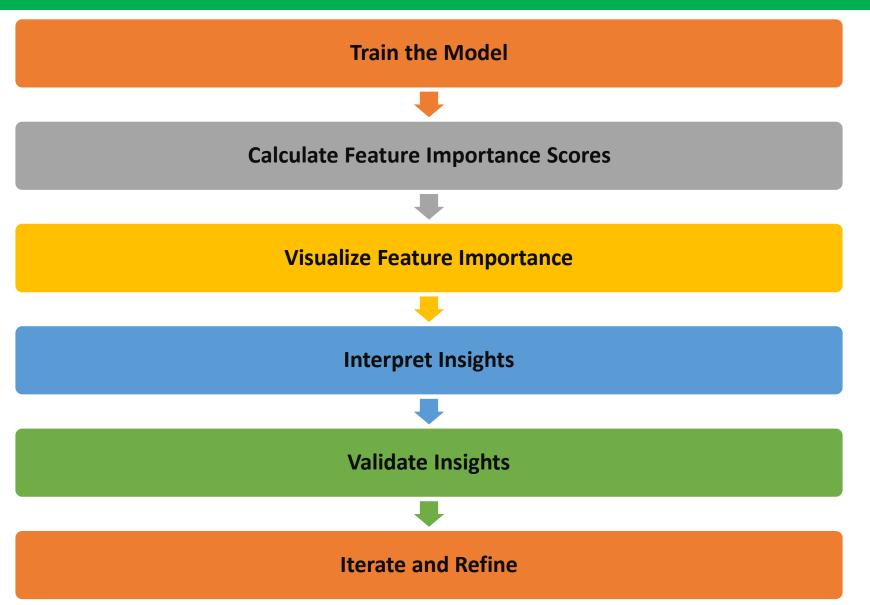
Feature engineering

Selecting subsets of features

Refining the modeling approach





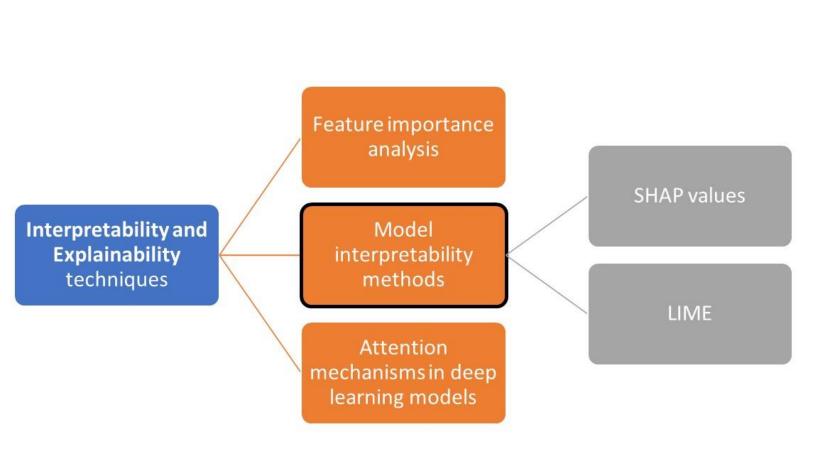






What is next?

What are Model Interpretability Methods to consider the interpretability and explainability of the selected models









Enrichmentors